Information Theory, Spring 2018	Iftach Haitner

Problem set 5

June 6, 2018	Due: June 21

- Please submit the handout in class, or email the grader (quefumas at gmail.com).
- Write clearly and shortly using sub-claims if needed. The emphasize in most questions is on the proofs (no much point is writing a “solution" w/o proving its correctness)
- For Latex users, a solution example can be found in the course web site.
- It is allowed to work in (small) groups, but please write the id list of your partners in the solution file, and each student should write his solution by *himself* (joint effort is only allowed in the “thinking phase")

1. (Coupling). Coupling is very useful tool in upper-bounding the statistical distance between two distributions. Here you are asked to prove and use a simple coupling lemma.

   (a) Prove that for pair of random variables (X, Y), it holds that $\mathrm{SD}(X, Y) \leq \Pr[X \neq Y]$. Is this bound tight?

   (b) Let P denote the the end point of n-step uniform random walk on $\mathbb{Z}$: start from 0, and at each step, move right with probability $1/2$ and Left otherwise.

   Let Q be the the end point of n-step δ-biased random walk on $\mathbb{Z}$: start from 0, and at each step, move right with probability $1/2 + \delta$ and Left otherwise.

   Use (a) to bound the statistical distance between P and Q.

2. (Bound on key size for almost perfect encryption)

   Let (E, D) be a perfectly correct encryption scheme for messages of length n and keys of length ℓ. Let $K \leftarrow \{0,1\}^\ell$. For each of the following cases find the best lower bound for ℓ.

   (a) $D(\mathsf{E}_K(m_0)||\mathsf{E}_K(m_1)) \leq \varepsilon$ for any $m_0, m_1 \in \{0,1\}^n$.

   (b) $\mathsf{SD}(\mathsf{E}_K(m_0), \mathsf{E}_K(m_1)) \leq \varepsilon$ for any $m_0, m_1 \in \{0,1\}^n$.

3. (Prediction to distinguishing) In class we showed that unpredictability implies indistinguishablity, here we prove that indistinguishablity implies unpredictability.

   (a) Let (X, Z) be a pair of random variables over $\{0,1\}^n \times \{0,1\}$. Let P be an s-size circuit such that
   $$\Pr[\mathsf{A}(Z) = X] \geq \frac{1}{2} + \varepsilon$$
   Prove there exists a circuit D of size s' not much larger than s, such that
   $$\Pr[\mathsf{D}(Z, X) = 1] - \Pr[\mathsf{B}(Z, U_1) = 1] \geq \varepsilon(n)$$
   where U_1 is uniformly distributed over $\{0,1\}$ (independently, of (X, Z).

   (b) Use (a) to show that if b is *not* an hardcore predicate of $f\colon \{0,1\}^n \mapsto \{0,1\}^n$ for s-size predictors, then $(f(U_n), b(U_n))$ is computationally *distinguishable* from $(f(U_n), U_1)$ by s' distinguisher, for s' not much smaller than s.

4. Let $f\colon \{0,1\}^n \mapsto \{0,1\}^n$ be (s,ε)-OWF, and let $\mathcal{H} = \{h\colon \{0,1\}^n \mapsto \{0,1\}^n\}$ be a pair-wise independent function family. Define g over $\{0,1\}^n \times \{0,1\}^n \times \mathcal{H} \times [n]$ by $g(x, r, h, i) = (f(x), r, h, h(x)_{1,\ldots,i}, b(x, r))$, for b being the Goldreich-Levin hardcore predicate (i.e., $b(x, r) = \langle x, r\rangle_2$). Find good as you can vales for s' and ε' such that $g(U_{2n}, H, I)$ has (s', ε')-entropy $H(g(U_{2n}, H, I)) + \frac{1}{2n}$, for $H \leftarrow \mathcal{H}$ and $I \leftarrow [n]$. You can assume that $\mathcal{H}$ can be sampled and evaluated by a size n circuit.